

Security analysis of decoy state quantum key distribution incorporating finite statistics

Jun Hasegawa,^{1,2} Masahito Hayashi,¹ Tohya Hiroshima,^{1,3} and Akihisa Tomita^{1,3}

¹ *Quantum Computation and Information Project, ERATO-SORST, Japan Science and Technology Agency,
Daini Hongo White Building 201, 5-28-3 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

² *Department of Computer Science, Graduate School of Information Science and Technology,
the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

³ *Nanoelectronics Research Laboratories, NEC Corporation, 34 Miyukigaoka, Tsukuba 305-8501, Japan*

Decoy state method quantum key distribution (QKD) is one of the promising practical solutions to BB84 QKD with coherent light pulses. In the real world, however, statistical fluctuations with the finite code length cannot be negligible, and the securities of theoretical and experimental researches of the decoy method state QKD so far are based on the asymptotic GLLP's formula which guarantees only that the limit of eavesdropper's information becomes zero as the code length approaches infinity. In this paper, we propose a substantially improved decoy state QKD in the framework of the finite code length and derive the upper bound of eavesdropper's information in the finite code length decoy state QKD with arbitrary number of decoy states of different intensities incorporating the finite statistics. We also show the performance of our decoy QKD and optimal values of parameters by numerical simulation.

PACS numbers: 03.67.Dd, 03.67.Hk, 03.67.-a

I. INTRODUCTION

Quantum key distribution (QKD) was originally proposed by Bennett and Brassard in 1984 [1] as a protocol, by which two parties, Alice and Bob, share secret keys by using a quantum communication channel as well as a public classical channel [2]. A remarkable feature is its unconditional security [3, 4, 5]; it is guaranteed by the fundamental laws of quantum mechanics and thereby QKD provides the unconditionally secure communication system. This is a triumph of quantum mechanics and quantum information science [6, 7] over the conventional cryptographic systems. In the practical setting of optical communication, however, it is the almost only option to substitute qubits in the original BB84 QKD protocol with heavily attenuated laser pulses because the perfect single photon emitting devices are not available in the current technology. Such laser pulses - the phase randomized weak coherent states - contains inevitably the multi-photon states at small but finite probability, which gives a malicious eavesdropper (Eve) a chance to obtain some amount of information on the shared keys by a photon-number-splitting attack [8]. Gottesman-Lo-Lütkenhaus-Preiskill (GLLP) showed, however, that it is still possible to obtain unconditionally secret key by BB84 protocol with such imperfect light sources, although the key generation rate and distances are very limited [9].

The recently proposed decoy state method [10, 11, 12, 13] is one of the promising practical solutions to BB84 QKD with coherent light pulses, in which several coherent state pulses with different intensities are used. Such optical pulses with different intensities have different photon number statistics. This simple fact equips Alice and Bob with a countermeasure against Eve. The original idea of the decoy state QKD is due to Hwang [10]. So far, several experimental demonstrations of decoy state QKD have been reported [14, 15, 16, 17, 18]. In most cases, the security analysis is based on the GLLP's asymptotic arguments, whereas, in the practical setting, the code length is finite so that the asymptotic argument is no longer valid and the *unconditional security* is actually not

guaranteed any more. The security analysis of QKD with the finite code size must incorporate the statistical fluctuations of the observed quantities [19]. Although several authors [11, 13, 15, 20] have considered the influence of statistical fluctuations on the decoy state QKD with finite code length, what all of them have done is limited to the re-adjustment of parameters of the asymptotic GLLP's formula. Such an *ad hoc* treatment cannot be justified to claim the unconditional security because the asymptotic GLLP's formula for the secure key generation rate provides us little knowledge about the eavesdropper's information when the finite code length is given. Therefore incorporating statistical fluctuations to the parameters of the asymptotic GLLP's formula cannot guarantee the security of the QKD protocols with the finite code length even if values of these parameters are exactly known. Suppose that the asymptotic rate of sacrifice bits needed for the secure final private key is R . When the code length is n , assigning nR to the number of sacrifice bits cannot ensure how secure the final key is at all. Because the asymptotic argument can only guarantee that the limit of eavesdropper's information becomes zero when the rate of sacrifice bits is greater than R , and without the speed of the convergence the increasing amount of sacrifice bits from nR , which is needed for the secure key with the code length n , cannot be estimated. Thus we must consider statistical fluctuations to eavesdropper's information formula with the finite code length n . Several upper bounds of eavesdropper's information with n have been provided [3, 5, 21], and especially Hayashi's formula is simple and gives better key generation rate than the others, some parameters of which cannot be directly obtained from observed quantities and are needed for estimation incorporating statistical fluctuations.

In this paper, we propose a substantially improved decoy state QKD in the framework of finite code length [21] by using the convex expansion formulas of weak coherent states [11, 21] and derive an eavesdropper's information formula in the finite code length decoy state QKD with arbitrary number of decoy states of different intensities incorporating

the finite statistics. We also show the performance of our decoy QKD and optimal values of parameters by numerical simulation.

The rest of this paper is organized as follows: In Section II, we begin by describing our decoy method QKD protocol. In Section III, we show Eve's information considering dark counts. We next explain how to estimate Eve's information in Section IV and random variables for describing the system in Section V, followed by the estimation incorporating statistical fluctuations in Section VI. We then demonstrate our numerical results of decoy method QKD protocol in Section VIII and finally summarize our results and discuss future work in section IX.

II. PROTOCOL

First of all, we describe our protocol [21]. We fix the size N of our code, the number N' of sent pulses, the maximum number \bar{N} and the minimum number \underline{N} of the size of a final key. We use $k+1$ different intensities or mean photon numbers $\mu_0 = 0 < \mu_1 < \dots < \mu_k$ including vacuum (μ_0) for the optical pulses. Two conjugate bases (+ and \times) are treated separately so that $2k+1$ different pulses are involved in total. The vacuum state ($i = 0$) is sent at the probability \bar{p}_0 and the μ_i pulse with \times (+) basis is sent at the probability \bar{p}_i (\bar{p}_{i+k}) ($i = 1, \dots, k$). The pulse with the intensity μ_{i_0} (μ_{i_0+k}) (the signal pulse) is used to distill a final secret key and the remainings (decoy pulses) are used just for estimation of Eve's attacks and/or the noise characteristics of quantum channel.

Before running the protocol, the probability p_D of dark counts in the detector and the other (basis-dependent) system error probability p_S (\bar{p}_S) of the \times (+) basis are measured in advance. The probability p_S or \bar{p}_S is the probability of errors other than the transmission errors, that is, the error probability for the noiseless channel. We assume that the detector is a threshold detector and the efficiency of the detector is independent of measurement bases [21].

The protocol is as follows. Alice randomly sends Bob a sequence of optical pulses of $k+1$ different intensities with randomly chosen basis. After that, Bob performs a measurement in one of the two bases and they compare bases and keep the pulses with the common basis by communicating via public channel. The number of sending pulses, received pulses, and pulses of the common basis are denoted by, respectively, A_i , C_i , and E_i ($i = 0, \dots, 2k$). Note that $\sum_{i=0}^{2k+1} A_i = N'$. The E_i bit string of i th kind of pulse contains error bits, which will be detected by checking a portion of the bits (check bits). To prepare check bits, they firstly perform the random permutation on E_{i_0} and E_{i_0+k} bit strings by sharing common random numbers via public channel. Then, for $i = i_0$ and $i = i_0 + k$, the first N bit string is used as the raw key and the remaining $E_{i_0} - N$ and $E_{i_0+k} - N$ bit string are used as the check bits, while the whole E_i bits are used as check bits for $i \neq i_0, i_0 + k$. (If $E_{i_0} \leq N$ or $E_{i_0+k} \leq N$, then the protocol is aborted.) The number of detected errors of i th kind of pulse is denoted by H_i ($i = 1, \dots, 2k$). From these quantities, they can evaluate the size of the final key guaranteeing the unconditional security.

If the evaluated final key size is not positive, the protocol is aborted again. The size of final secret key of + basis is computed as

$$N_{final} := N\eta\left(\frac{H_{i_0+k}}{E_{i_0+k} - N}\right) - m(\mathcal{D}_i, \mathcal{D}_e), \quad (1)$$

and that of \times basis is

$$\hat{N}_{final} := N\eta\left(\frac{H_{i_0}}{E_{i_0} - N}\right) - \tilde{m}(\tilde{\mathcal{D}}_i, \mathcal{D}_e), \quad (2)$$

where $\eta(\cdot)$ denotes the error correcting coding rate and $m(\mathcal{D}_i, \mathcal{D}_e)$ and $\tilde{m}(\tilde{\mathcal{D}}_i, \mathcal{D}_e)$ represents the size of privacy amplification. Here we abbreviate the initial data $(\mathbf{A}, \boldsymbol{\mu}, p_S, p_D)$, $(\mathbf{A}, \boldsymbol{\mu}, \bar{p}_S, \bar{p}_D)$ and the observed data $(\mathbf{C}, \mathbf{E}, \mathbf{H})$ to \mathcal{D}_i , $\tilde{\mathcal{D}}_i$, and \mathcal{D}_e , respectively, and $\mathbf{A} = (A_1, \dots, A_{2k})$, etc. If $N_{final} < \underline{N}$ or $\hat{N}_{final} < \underline{N}$, they abort the protocol and go back to the first step. Furthermore, if $\bar{N} < N_{final}$ ($\bar{N} < \hat{N}_{final}$), they replace $m(\mathcal{D}_i, \mathcal{D}_e)$ [$\tilde{m}(\tilde{\mathcal{D}}_i, \mathcal{D}_e)$] by $N\eta(H_{i_0+k}/(E_{i_0+k} - N)) - \bar{N}$ [$N\eta(H_{i_0}/(E_{i_0} - N)) - \bar{N}$]. Finally, they are left with N bits error correction followed by privacy amplification to share the N_{final} (\hat{N}_{final}) bit secret key of + (\times) basis.

The error correction is performed as follows. Suppose that Alice and Bob have, respectively, the random number sequences X and X' of N bits, which contain some errors. The task is to distill the common random number sequence of $l+m$ bits with negligible errors. In the forward error correction, they share $N \times (l+m)$ binary matrix M_e . Alice generates other $l+m$ bits random number Z , and sends a bit sequence $M_e Z + X$ to Bob. Then, Bob applies the decoding of the code M_e to the bit sequence $M_e Z + X - X'$ to extract Z . On the other hand, in the reverse error correction, Bob generates the random number sequence Z of $l+m$ bits, and sends $M_e Z + X'$ to Alice. Then, Alice applies the decoding code M_e to the bit sequence $M_e Z + X' - X$ to extract Z . The error correction here corresponds to a part of the twirling operation so that their channel can be regarded as a Pauli channel from Alice (Bob) to Bob (Alice) in the forward (reverse) error correction [21].

In the privacy amplification, Alice and Bob share the final secret key of l bits from Z of $l+m$ bits. More precisely, they first generate the same $l \times (l+m)$ binary matrix M_p with

$$\text{Prob}\{Z \in \text{Im}M_p^T\} \leq 2^{-m} \quad (3)$$

for any non-zero $l+m$ bit sequence Z . Subsequently, they generate the bit sequence $M_p Z$ of l bits from Z of $l+m$ bits.

Combining the error correction and the privacy amplification described above, the sequel of it is that Alice sends information by the code $\text{Im}M_e/M_e(\text{Ker}M_p)$.

III. GENERAL UPPER BOUND FOR EVE'S INFORMATION ON FINAL KEY

In this section, we give an upper bound for the leakage information on the final key, which lays the foundation of the security analysis in Sec. VI [21]. Here, we confine ourselves

to the discussion on the final key with \times basis. Eve's attack can be described by the conditional distribution \mathcal{P} of the Pauli action on the input state. Hence, the average of Eve's information with respect to the final key is closely related to the error probability $P_{ph,min,x|M_p,D_e,POS}^{\mathcal{P}}$ with the minimum distance decoding when information is sent with \times basis and the code $\text{Im}M_e/M_e(\text{Ker}M_p)$, where POS is a random variable for the arrangement of different intensities and the position of check bits, and $x \Rightarrow (\leftarrow)$ refers to the forward (reverse) error correction. The average $I_{E,av,x}^{\mathcal{P}} = \mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} I_{E,x|M_p,D_e,POS}^{\mathcal{P}}$ of Eve's information $I_{E,x|M_p,D_e,POS}^{\mathcal{P}}$ is evaluated in terms of $P_{ph,av,x}^{\mathcal{P}} = \mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} P_{ph,min,x|M_p,D_e,POS}^{\mathcal{P}}$ as

$$I_{E,av,x}^{\mathcal{P}} \leq P_{ph,av,x}^{\mathcal{P}} (1 + \bar{N} - \log P_{ph,av,x}^{\mathcal{P}}). \quad (4)$$

Since the stochastic behavior of the random variables \mathcal{D}_e depends on the conditional distribution \mathcal{P} , we denote the operation of taking the expectation with respect to $M_p, \mathcal{D}_e, \text{POS}$ by $\mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}}$. Denoting Eve's state with respect to the final key $[Z] = M_p Z$ by $\rho_{[Z]}^{E,x}$, and its average state by $\bar{\rho}^{E,x}$, we obtain the following inequalities.

$$\mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} \min_{[Z] \neq [Z']} F(\rho_{[Z]}^{E,x}, \rho_{[Z']}^{E,x}) \geq 1 - 2P_{ph,av,x}^{\mathcal{P}}, \quad (5)$$

$$\mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} \max_{[Z] \neq [Z']} \|\rho_{[Z]}^{E,x} - \rho_{[Z']}^{E,x}\|_1 \leq 4P_{ph,av,x}^{\mathcal{P}}, \quad (6)$$

$$\mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} \min_{[Z]} F(\rho_{[Z]}^{E,x}, \bar{\rho}^{E,x}) \geq 1 - 2P_{ph,av,x}^{\mathcal{P}}, \quad (7)$$

and

$$\mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} \max_{[Z]} \|\rho_{[Z]}^{E,x} - \bar{\rho}^{E,x}\|_1 \leq 4P_{ph,av,x}^{\mathcal{P}}. \quad (8)$$

Here, we have omitted the dependence of $\rho_{[Z]}^{E,x}$ on M_p, \mathcal{D}_e , and POS. Next, let $P_{succ,x|M_p}^{\mathcal{P}}$ be the probability that Eve acquires perfectly information on the final key when she performs the optimal measurement after the privacy amplification. Then,

$$\begin{aligned} & \mathbb{E}_{M_p,D_e,POS}^{\mathcal{P}} P_{succ,x|M_p}^{\mathcal{P}} \\ & \leq \left(\sqrt{P_{ph,av,x}^{\mathcal{P}}} \sqrt{1 - 2^{-\bar{N}}} + \sqrt{1 - P_{ph,av,x}^{\mathcal{P}}} \sqrt{2^{-\bar{N}}} \right)^2 \end{aligned} \quad (9)$$

Here, we have again omitted the dependence of $P_{succ,x|M_p}^{\mathcal{P}}$ on M_p, \mathcal{D}_e , and POS. Now, it is evident that the evaluation of $P_{ph,av,x}^{\mathcal{P}}$ plays an essential role in the security analysis.

We start by grouping detected pulses into six parts according to which state (vacuum, single photon, or multi-photon) is actually sent by Alice and whether or not the detection is normal, i.e., it is *not* due to the dark counts. We define J^i (J^{3+i}) as the number of pulses detected normally (by dark counts) under the condition that the state sent is vacuum ($i = 0$), single photon ($i = 1$), or multi-photon ($i = 2$) states. For example, J^3 represents the number of pulses detected by dark counts when the state sent by Alice is the vacuum. We regard the simultaneous event of a dark count and a normal count as a dark count. This is because the collision of both photons causes

the information loss of the normal count. Let t be the number of pulses or bits with transmission (phase) error in \times basis among J^1 bits. This is also a random variable. Then, by denoting the expectation with respect to the random variables t and $\mathbf{J} = (J^0, \dots, J^5)$ by $\mathbb{E}_{t,\mathbf{J}}^{\mathcal{P}}$, $P_{ph,av,x}^{\mathcal{P}}$ can be evaluated as

$$\begin{aligned} P_{ph,av,\rightarrow}^{\mathcal{P}} & \leq \mathbb{E}_{t,\mathbf{J},D_e,POS}^{\mathcal{P}} 2^{-[m(\mathcal{D}_i, \mathcal{D}_e) - J^1 \bar{h}(t/J^1) - J^2 - J^4 - J^5]_+} \\ & = \mathbb{E}_{t,\mathbf{J},D_e,POS}^{\mathcal{P}} 2^{-[m(\mathcal{D}_i, \mathcal{D}_e) - N + J^1 (1 - \bar{h}(t/J^1)) + J^0 + J^3]_+}, \end{aligned} \quad (10)$$

and

$$\begin{aligned} P_{ph,av,\leftarrow}^{\mathcal{P}} & \leq \mathbb{E}_{t,\mathbf{J},D_e,POS}^{\mathcal{P}} 2^{-[m(\mathcal{D}_i, \mathcal{D}_e) - J^1 \bar{h}(t/J^1) - J^0 - J^2]_+} \\ & = \mathbb{E}_{t,\mathbf{J},D_e,POS}^{\mathcal{P}} 2^{-[m(\mathcal{D}_i, \mathcal{D}_e) - N + J^1 (1 - \bar{h}(t/J^1)) + J^3 + J^4 + J^5]_+}, \end{aligned} \quad (11)$$

where $[z]_+ = \max\{0, z\}$ and $\bar{h}(x)$ is defined by

$$\bar{h}(x) := \begin{cases} -x \log_2 x - (1-x) \log_2 (1-x) & \text{if } x \in [0, 1/2] \\ 1 & \text{if } x \in (1/2, 1]. \end{cases} \quad (12)$$

In the actual system, the random variables t and \mathbf{J} cannot be identified exactly. They are estimated from the measured values \mathcal{D}_e , by which the size of sacrifice bits $m(\mathcal{D}_i, \mathcal{D}_e)$ is determined. It is of crucial importance to determine $m(\mathcal{D}_i, \mathcal{D}_e)$ such that the average error probability $P_{ph,av,x}^{\mathcal{P}}$ is less than a given value for any attack \mathcal{P} . The statistical fluctuation of \mathcal{D}_e is properly taken into account in the computation of $m(\mathcal{D}_i, \mathcal{D}_e)$. As for the attack \mathcal{P} , it is sufficient to treat the extremal points, in which these random variables can be described by the combination of multi-hypergeometric distributions. All random variables concern our problem are listed in Sec. V.

IV. EVE'S STRATEGY AND ITS ESTIMATION

Suppose that Eve can distinguish the different number states. A naïve way to describe Eve's attacks is to associate each number state with the corresponding parameters describing Eve's attacks. This is, however, a formidable task because the infinite number of unknown parameters are involved. In order to avoid such a difficulty, one of the authors [22] introduced a convex expansion of the phase-randomized coherent state $\sum_{n=0}^{\infty} e^{-\mu} \mu^n |n\rangle \langle n| / n!$ in terms of vacuum, single-photon, and multi-photon states. Here, we define the multi-photon states σ_l ($l = 2, \dots, k+1$) as

$$\sigma_l := \frac{1}{\Omega_l} \sum_{n=l}^{\infty} \frac{\gamma_{l,n}}{n!} |n\rangle \langle n|, \quad (13)$$

where

$$\Omega_l := \sum_{n=l}^{\infty} \frac{\gamma_{l,n}}{n!}, \quad (14)$$

and

$$\gamma_{l,n} := \sum_{j=1}^{l-1} \frac{\mu_j^{n-2}}{\prod_{t=1, t \neq j}^{l-1} (\mu_j - \mu_t)}, \quad (15)$$

with $\mu_1 < \mu_2 < \dots < \mu_k$. Note that σ_l [Eq. (13)] are *bona fide* states, i.e., $\sigma_l \geq 0$ and $\text{Tr}\sigma_l = 1$. It is easy to see that the state $\sum_{n=0}^{\infty} e^{-\mu} \mu^n |n\rangle\langle n|/n!$ can be expressed as a convex combination of $|0\rangle\langle 0|$, $|1\rangle\langle 1|$, and σ_l ;

$$e^{-\mu_i} \sum_{n=0}^{\infty} \frac{\mu_i^n}{n!} |n\rangle\langle n| = e^{-\mu_i} \left(|0\rangle\langle 0| + \mu_i |1\rangle\langle 1| + \sum_{n=2}^{i+1} \mu_i^2 \prod_{t=1}^{n-2} (\mu_i - \mu_t) \Omega_n \sigma_n \right). \quad (16)$$

Here, coefficients

$$e^{-\mu_i} \mu_i^2 \prod_{t=1}^{n-2} (\mu_i - \mu_t) \Omega_n$$

are positive.

Now, we adopt the worst case scenario on Eve's attacks. Namely, we assume that Eve can distinguish vacuum state ($\rho_0 = |0\rangle\langle 0|$), single photon state ($\rho_1 = |1\rangle\langle 1|$), multi-photon states with \times basis $\rho_j = \sigma_j$ ($j = 2, \dots, k+1$) and those with $+$ basis $\rho_{k+j} = \sigma_j$ ($j = k+2, \dots, 2k+1$). The number of emitted j th state ρ_j ($j = 0, \dots, 2k+1$) is denoted by B^j ; $\mathbf{B} = (B^j)$. According to the values of \mathbf{B} , Eve can do the following attacks; Eve tricks Bob into detecting j th state with

ratio $q^j(\mathbf{B})$ [23] and causes phase errors with ratio $r^j(\mathbf{B})$ for j th state ($j = 1, 2, \dots, k+1$) and bit errors with ratio $\tilde{r}^j(\mathbf{B})$ for j th state ($j = 1, k+2, \dots, 2k+1$). The quantities $r^j(\mathbf{B})$ and $\tilde{r}^j(\mathbf{B})$ describe the transmission errors. In the following, we focus on the final secret key of \times basis and write q^j and r^j (\tilde{r}^j) instead of $q^j(\mathbf{B})$ and $r^j(\mathbf{B})$ [$\tilde{r}^j(\mathbf{B})$]. Note that q^j is the rate of detection with the exclusion of dark counts. Since the state ρ_j with the \times basis is, in general, different from that with the $+$ basis, the parameters q^j do not necessarily coincide with q^{j+k} ($j = 1, \dots, k$). The generating probability of each state can be described by the matrix $(P_{k,i}^j)_{i=0, \dots, 2k, j=0, \dots, 2k+1}$ defined by

$$P_k := \begin{pmatrix} 1 & 0 & 0 & 0 \\ Y & Z & X & 0 \\ Y & Z & 0 & X \end{pmatrix} \quad (17)$$

where Y and Z are k -dimensional vectors such that $Y_i = e^{-\mu_i}$ and $Z_i = \mu_i e^{-\mu_i}$ and the $k \times k$ matrix X is given by

$$X_i^j := \begin{cases} \mu_i^2 \prod_{t=1}^{j-1} (\mu_i - \mu_t) e^{-\mu_i} \Omega_{j+1} & j = 1, \dots, i \\ 0 & j = i+1, \dots, k, \end{cases} \quad (18)$$

for $i = 1, \dots, k$. For example, the matrix P_k for $k = 3$ is given by

$$P_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ e^{-\mu_1} & \mu_1 e^{-\mu_1} & e^{-\mu_1} \mu_1^2 \Omega_2 & 0 & 0 & 0 & 0 & 0 \\ e^{-\mu_2} & \mu_2 e^{-\mu_2} & e^{-\mu_2} \mu_2^2 \Omega_2 & e^{-\mu_2} \mu_2^2 (\mu_2 - \mu_1) \Omega_3 & 0 & 0 & 0 & 0 \\ e^{-\mu_3} & \mu_3 e^{-\mu_3} & e^{-\mu_3} \mu_3^2 \Omega_2 & e^{-\mu_3} \mu_3^2 (\mu_3 - \mu_1) \Omega_3 & * & 0 & 0 & 0 \\ e^{-\mu_1} & \mu_1 e^{-\mu_1} & 0 & 0 & 0 & e^{-\mu_1} \mu_1^2 \Omega_2 & 0 & 0 \\ e^{-\mu_2} & \mu_2 e^{-\mu_2} & 0 & 0 & 0 & e^{-\mu_2} \mu_2^2 \Omega_2 & e^{-\mu_2} \mu_2^2 (\mu_2 - \mu_1) \Omega_3 & 0 \\ e^{-\mu_3} & \mu_3 e^{-\mu_3} & 0 & 0 & 0 & e^{-\mu_3} \mu_3^2 \Omega_2 & e^{-\mu_3} \mu_3^2 (\mu_3 - \mu_1) \Omega_3 & * \end{pmatrix}, \quad (19)$$

$$* = e^{-\mu_3} (\mu_3 - \mu_1) (\mu_3 - \mu_2) \Omega_4.$$

Now, the expectation of the detection probability $p_i = C_i/A_i$ is expressed as

$$\mathbb{E}(p_i) = \Xi_i^{(1)}(\mathbf{q}) := \sum_{j=0}^{2k+1} P_i^j q^j + p_D, \quad (20)$$

for $i = 0, \dots, 2k$ and the expectation of detected error probability is written as

$$\mathbb{E}(s_i p_i) = \Xi_i^{(2)}(\mathbf{q}, \mathbf{r}) := P_i^1 q^1 r^{1'} + \sum_{j=2}^{k+1} P_i^j q^j r^j + \frac{1}{2} (P_i^0 q^0 + p_D), \quad (21)$$

for $i = 1, 2, \dots, k$, where $s_i = H_i/E_i$ ($i \neq i_0$) and $s_{i_0} = H_{i_0}/(E_{i_0} - N)$. In Eq. (21), $r^{1'} = (1 - p_S) r^1 + p_S (1 - r^1)$. Since we cannot uniquely determine the parameters q^j and r^j from Eqs. (20) and (21) even when $\mathbb{E}(p_i) = p_i$ and $\mathbb{E}(s_i p_i) = s_i p_i$, we fix $\frac{q^{k+1} + q^{2k+1}}{\sqrt{2}}$ and r^{k+1} as $x \in [0, \sqrt{2}(1 - p_D)]$ and $y \in [0, 1]$

to obtain

$$\hat{q}_x^j = \begin{cases} p_0 - p_D & j = 0 \\ \xi^j & j = 1, \dots, k, k+2, \dots, 2k \\ \frac{1}{\sqrt{2}} (x + \xi^{k+1}) & j = k+1 \\ \frac{1}{\sqrt{2}} (x - \xi^{k+1}) & j = 2k+1, \end{cases} \quad (22)$$

and

$$\hat{r}_{x,y}^j = \begin{cases} \frac{1}{1-2p_S} (\xi^1 - p_S) & j = 1 \\ \xi^j & j = 2, \dots, k \\ y & j = k+1, \end{cases} \quad (23)$$

with

$$\xi^j := \frac{1}{\hat{q}_x^j} \sum_{i=1}^k (P_{k \times k}^{-1})_{ij}^i \times \left\{ s_i p_i - \frac{1}{2} [P_i^0 (p_0 - p_D) + p_D] - P_i^{k+1} \hat{q}_x^{k+1} y \right\}, \quad (24)$$

where $X_{l \times l}$ presents the $l \times l$ submatrix $(X_{i,j})_{1 \leq i,j \leq l}$ for a given rectangular matrix $X = (X_{i,j})_{0 \leq i \leq a, 0 \leq j \leq b}$, and ξ^j ($j = 1, \dots, 2k$) in Eq. (22) is defined as

$$\xi^j := \sum_{i=1}^{2k} (\bar{P}_{2k \times 2k}^{-1})_{ij}^i \left[p_i - p_D - \bar{P}_i^0 (p_0 - p_D) - x \bar{P}_i^{2k+1} \right], \quad (25)$$

where

$$\bar{P}_i^j := \sum_{j'=0}^{2k+1} P_i^{j'} Q^{j',j}, \quad (26)$$

with

$$Q^{j',j} := \begin{cases} -\frac{1}{\sqrt{2}} & j' = k+1, j = 2k+1 \\ \frac{1}{\sqrt{2}} & j' = 2k+1, j = 2k+1 \\ \frac{1}{\sqrt{2}} & j' = 2k+1 \text{ or } k+1, j = k+1 \\ \delta^{j',j} & \text{Otherwise} \end{cases} \quad (27)$$

$$\begin{aligned} \left(\mathbf{q}_{x,y,(B_i^j)}^{ML}, \mathbf{r}_{x,y,(B_i^j)}^{ML} \right) = & \underset{\mathbf{q}, \mathbf{r}: q^{k+1} + q^{2k+1} = \sqrt{2}x, r^{k+1} = y}{\text{argmax}} \left\{ \sum_{i=0}^{2k} \left[C_i \log \Xi_i^{(1)}(\mathbf{q}) + (A_i - C_i) \log(1 - \Xi_i^{(1)}(\mathbf{q})) \right] \right. \\ & \left. + \sum_{i=1}^k H_i \log \frac{\Xi_i^{(2)}(\mathbf{q}, \mathbf{r})}{\Xi_i^{(1)}(\mathbf{q})} + \sum_{i=1}^k (E_i - \delta_{i,i_0} N - H_i) \log \left(1 - \frac{\Xi_i^{(2)}(\mathbf{q}, \mathbf{r})}{\Xi_i^{(1)}(\mathbf{q})} \right) \right\}, \end{aligned} \quad (28)$$

with $0 \leq q^j \leq 1 - p_D$ and $0 \leq r^j \leq 1$. If the above linear estimators Eqs. (22) and (23) are in the range $0 \leq \hat{q}_x^j \leq 1 - p_D$ and $0 \leq \hat{r}_{x,y}^j \leq 1$, then they coincide with the corresponding maximum likelihood estimators.

V. RANDOM VARIABLES DESCRIBING THE SYSTEM

To compute the size of final secret key with the finite statistics due to the finite code length, several stochastic variables must be incorporated properly. In this section, we describe random variables with their means and (co)variances, which are used in the computation of the sacrifice key size (Sec. VII).

Firstly, we define B_i^j as the number of the emitted state ρ_j given that the i th kind of pulse is sent ($i = 0, \dots, 2k$; $j = 0, \dots, 2k+1$). They obey the following multi-nomial distribution.

$$P(B_i^0, \dots, B_i^{2k+1}) = (P_i^0)^{B_i^0} \dots (P_i^{2k+1})^{B_i^{2k+1}} \frac{A_i!}{B_i^0! \dots B_i^{2k+1}!}. \quad (29)$$

Note that $A_i = \sum_{j=0}^{2k+1} B_i^j$ and $B^j = \sum_{i=0}^{2k} B_i^j$. The mean of B_i^j is $P_i^j A_i$.

Next, we define C_i^j as the number of i th kind of pulse detected normally under the condition that the emitted states is ρ_j ($i = 0, \dots, 2k$; $j = 0, \dots, 2k+1$). Note that the dark counts are not included in the detection events. The contribution of dark counts is expressed as C_i^{-1} ($j = -1$). These stochastic

When the true values q^j and r^j are close to zero, say, the linear estimators given above often take on negative values due to statistical fluctuations. The maximal likelihood estimation provides an alternate solution free from such a drawback, which is given by

variables obey

$$P(C_0^j, C_1^j, \dots, C_{2k}^j) = \frac{\binom{B_0^j}{C_0^j} \binom{B_1^j}{C_1^j} \dots \binom{B_{2k}^j}{C_{2k}^j}}{\binom{\sum_{i=0}^{2k} B_i^j}{q^j \sum_{i=0}^{2k} B_i^j}}, \quad (30)$$

for $j \neq -1$ and

$$P(C_i^{-1}) = p_D^{C_i^{-1}} (1 - p_D)^{A_i - C_i^{-1}} \binom{A_i}{C_i^{-1}}. \quad (31)$$

Note that $\sum_{i=0}^{2k} C_i^j = q^j \sum_{i=0}^{2k} B_i^j$ and $C_i = \sum_{j=-1}^{2k+1} C_i^j$. The means $\mathbb{E}C_i^j$ are given by $q^j B_i^j$ for $j \neq -1$ and $p_D A_i$ for $j = -1$. Deviations $\Delta' C_i^j = C_i^j - \mathbb{E}C_i^j$ satisfy

$$\begin{aligned} \mathbb{E} \Delta' C_i^j \Delta' C_{i'}^{j'} &= q^j (1 - q^j) \left(\delta_{i,i'} B_i^j - \frac{B_i^j B_{i'}^{j'}}{\sum_{i=0}^{2k} B_i^j} \right) \\ &\cong q^j (1 - q^j) \left(\delta_{i,i'} P_i^j A_i - \frac{P_i^j A_i P_{i'}^{j'} A_{i'}}{\sum_{i=0}^{2k} P_i^j A_i} \right), \end{aligned} \quad (32)$$

for $j \neq -1$ and

$$\mathbb{E}(\Delta' C_i^{-1})^2 = p_D (1 - p_D) A_i. \quad (33)$$

Other covariances are zero.

Since Bob measures the received pulses with randomly chosen basis, the measuring basis coincides with the basis of ρ_j

with probability $1/2$ [24]. Therefore, defining the numbers of common basis pulses among C_i^j by $E_i^j = \frac{1}{2}C_i^j + \Delta'E_i^j$, they obey the following binomial distribution

$$P(E_i^j) = \binom{\frac{1}{2}C_i^j}{E_i^j}, \quad (34)$$

for $i = 0, 1, \dots, 2k$ and $j = -1, \dots, 2k + 1$, and the nonzero covariances of $\Delta'E_i^j$ are computed as

$$\mathbb{E}(\Delta'E_i^j)^2 = \frac{1}{k+1}C_i^j \cong \frac{q^j P_i^j A_i}{k+1}, \quad (35)$$

for $j \neq -1$ and

$$\mathbb{E}(\Delta'E_i^{-1})^2 = \frac{1}{k+1}C_i^{-1} \cong \frac{p_D A_i}{k+1}. \quad (36)$$

Now we define the following quantities F_i^j ; $F_{i_0}^j$ ($F_{i_0+k}^j$) denotes the number of check bits with \times (+) basis within $E_{i_0} - N$ (N) bits given that the emitted state is ρ_j and $F_i^j = E_i^j$ for $i = 1, \dots, k, i \neq i_0$. For $i = i_0$ and $i = i_0 + k$, their distributions are, respectively, given by the following multi-hypergeometric distributions ($j = -1, 0, 1, \dots, 2k + 1$).

$$\begin{aligned} & P(F_{i_0}^{-1}, F_{i_0}^0, F_{i_0}^1, F_{i_0}^2, \dots, F_{i_0}^{k+1}) \\ &= \frac{\binom{E_{i_0}^{-1}}{F_{i_0}^{-1}} \binom{E_{i_0}^0}{F_{i_0}^0} \binom{E_{i_0}^1}{F_{i_0}^1} \binom{E_{i_0}^2}{F_{i_0}^2} \dots \binom{E_{i_0}^{k+1}}{F_{i_0}^{k+1}}}{\binom{E_{i_0} - N}{N}}, \end{aligned} \quad (37)$$

and

$$\begin{aligned} & P(F_{i_0+k}^{-1}, F_{i_0+k}^0, F_{i_0+k}^1, F_{i_0+k}^2, \dots, F_{i_0+k}^{2k+1}) \\ &= \frac{\binom{E_{i_0+k}^{-1}}{F_{i_0+k}^{-1}} \binom{E_{i_0+k}^0}{F_{i_0+k}^0} \binom{E_{i_0+k}^1}{F_{i_0+k}^1} \binom{E_{i_0+k}^2}{F_{i_0+k}^2} \dots \binom{E_{i_0+k}^{2k+1}}{F_{i_0+k}^{2k+1}}}{\binom{E_{i_0+k}}{N}}, \end{aligned} \quad (38)$$

where $E_i = \sum_{j=-1}^{2k+1} E_i^j$, $\sum_{j=-1}^{2k+1} F_{i_0}^j = E_{i_0} - N$, $\sum_{j=-1}^{2k+1} F_{i_0+k}^j = N$. Note that $E_{i_0}^{k+2} = \dots = E_{i_0}^{2k+1} = E_{i_0+k}^2 = \dots = E_{i_0+k}^{k+1} = 0$. It is easy to see that

$$\mathbb{E}F_{i_0}^j = \frac{E_{i_0} - N}{E_{i_0}} E_{i_0}^j, \quad (39)$$

and

$$\mathbb{E}F_{i_0+k}^j = \frac{N}{E_{i_0+k}} E_{i_0+k}^j, \quad (40)$$

for $j = -1, \dots, 2k + 1$. The nonzero covariances of deviation $\Delta'F_i^j$ are computed as

$$\mathbb{E}\Delta'F_i^j \Delta'F_i^{j'} = \frac{E_i^j}{E_i} \left(1 - \frac{N}{E_i}\right) \left(\delta^{jj'} - \frac{NE_i^{jj'}}{E_i}\right), \quad (41)$$

for $i = i_0$ ($j = -1, 0, 1, 2, \dots, k + 1$) and $i = i_0 + k$ ($j = -1, 0, 1, k + 2, \dots, 2k + 1$).

The errors occur among F_i^j check bits in \times basis with probability r^j ($j = 1, \dots, k + 1$). We define G_i^j as the number of pulses with transmission errors in \times basis among F_i^j pulses; $G_i^j = r^j F_i^j + \Delta'G_i^j$, which obey the following multi-hypergeometric distribution

$$\begin{aligned} & P(G_1^j, \dots, G_k^j, G_{i_0+k}^j) \\ &= \frac{\binom{F_1^j}{G_1^j} \dots \binom{F_k^j}{G_k^j} \binom{F_{i_0+k}^j}{G_{i_0+k}^j} \binom{\sum_{i'=0}^{2k} C_{i'} - \sum_{i'=1}^{2k} F_{i'}^j}{\sum_{i'=0}^{2k} C_{i'} - \sum_{i'=1}^{2k} G_{i'}^j}}{\binom{\sum_{i'=0}^{2k} C_{i'}}{\sum_{i'=1}^{2k} C_{i'}}}, \end{aligned} \quad (42)$$

for $i = 1, \dots, k, i_0 + k$ and $j = 1, 2, \dots, k + 1$. Note that $G_i^j = 0$ for $i \geq k + 1$ and $i \neq i_0 + k$ and that the system errors other than the transmission errors are *not* counted in the definition of G_i^j . The nonzero covariances of deviations $\Delta'G_i^j$ are computed as

$$\mathbb{E}\Delta'G_i^j \Delta'G_{i'}^{j'} = r^j (1 - r^j) \left(\delta_{i,i'} F_i^j - \frac{F_i^j F_{i'}^{j'}}{\sum_{i=0}^{2k} C_i^j} \right). \quad (43)$$

The number of detected errors in \times basis H_i is the sum of several contributions. The contribution of dark counts (vacuum state) is denoted by H_i^{-1} (H_i^0), which is the number of detected errors of ρ_{-1} (ρ_0) among F_i^{-1} (F_i^0) bits. Since the bits received by Bob are completely independent of the bits sent by Alice for $j = -1$ and 0 , the error probability is $1/2$ so that the probability distributions of random the random variables H_i^{-1} and H_i^0 are, respectively, given by

$$P(H_i^{-1}) = \left(\frac{1}{2}\right)^{F_i^{-1}} \binom{F_i^{-1}}{H_i^{-1}}, \quad (44)$$

and

$$P(H_i^0) = \left(\frac{1}{2}\right)^{F_i^0} \binom{F_i^0}{H_i^0}. \quad (45)$$

For the single photon state, the errors occurred within G_i^1 bits are recovered accidentally by the system errors other than transmission errors with probability p_S and that the errors occurred within $F_i^1 - G_i^1$ by the same cause contribute to the detected errors with probability p_S . Therefore, the detected errors of the single photon state are divided into two; H_i^1 and $H_i^{1'}$, whose probability distributions are, respectively, given by

$$P(H_i^1) = (1 - p_S)^{H_i^1} p_S^{G_i^1 - H_i^1} \binom{G_i^1}{H_i^1}, \quad (46)$$

and

$$P(H_i^{1'}) = p_S^{H_i^{1'}} (1 - p_S)^{F_i^1 - G_i^1 - H_i^{1'}} \binom{F_i^1 - G_i^1}{H_i^{1'}}. \quad (47)$$

The random variable H_i is then written as

$$\begin{aligned} H_i &= H_i^{-1} + H_i^0 + H_i^1 + H_i^{1'} + \sum_{j=2}^{k+1} G_i^j \\ &= \frac{1}{2}(F_i^{-1} + F_i^0) + (1 - p_S)G_i^1 + p_S(F_i^1 - G_i^1) \\ &\quad + \sum_{j=2}^{k+1} G_i^j + \Delta' H_i, \end{aligned} \quad (48)$$

where $\Delta' H_i$ is the deviation whose nonzero variances are given by

$$\mathbb{E}(\Delta' H_i)^2 = \frac{1}{k+1}(F_i^{-1} + F_i^0) + p_S(1 - p_S)F_i^1. \quad (49)$$

The numbers known by Alice and Bob are $C_i = \sum_{j=-1}^{2k+1} C_i^j$, $E_i = \sum_{j=-1}^{2k+1} E_i^j$ for $i = 0, 1, \dots, 2k$, and H_i for $i = 1, \dots, k$.

VI. COMPUTATION OF SACRIFICE KEY SIZE (REVERSE CASE)

In this section, we give a method to derive the size of sacrifice key of + basis in case of reverse error correction. According to the central limit theorem with respect to the multinomial and multi-hypergeometric distributions, we can assume safely that all random variables given in Sec. V obey normal distributions with the averages and the (co)variances given in Sec. V because the number of our samples is sufficiently large.

In the following argument, stochastic variables B_i^j are fixed. For now, we fix also $x = \frac{q^{k+1}(\mathbf{B}) + q^{2k+1}(\mathbf{B})}{\sqrt{2}}$ and $y = r^{k+1}(\mathbf{B})$. Applying the inequality (11), the quantity $P_{ph,av,\leftarrow}^P$ is bounded from above by the expectation of

$$2 \left[m(\mathcal{D}_i, \mathcal{D}_e) - N + F_{i_0+k}^1 \left(1 - \bar{h} \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right) + F_{i_0+k}^{-1} \right]_+.$$

Here, J^1 and t in (11) are, respectively, given by $F_{i_0+k}^1$ and $G_{i_0+k}^1$. The number of pulses detected by dark counts $J^3 + J^4 + J^5$ in (11) is simply expressed as $F_{i_0+k}^{-1}$. Now, we introduce a new function $\bar{h}_a(x)$ which is a slight modification of $\bar{h}(x)$:

$$\bar{h}_a(x) = \begin{cases} \bar{h}(x) & \text{if } x \geq a, \\ \bar{h}(a) + \bar{h}'(a)(x - a) & \text{if } x < a. \end{cases} \quad (50)$$

Owing to the convexity of $\bar{h}(x)$, we have $\bar{h}(x) \leq \bar{h}_a(x)$ to obtain

$$\begin{aligned} &2 \left[m(\mathcal{D}_i, \mathcal{D}_e) - N + F_{i_0+k}^1 \left(1 - \bar{h} \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right) + F_{i_0+k}^{-1} \right]_+ \\ &\leq 2 \left[m(\mathcal{D}_i, \mathcal{D}_e) - N + F_{i_0+k}^1 \left(1 - \bar{h}_a \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right) + F_{i_0+k}^{-1} \right]_+, \end{aligned} \quad (51)$$

which is used for an upper bound of the quantity $P_{ph,av,\leftarrow}^P$. Here, we estimate

$$\Theta(\mathcal{D}_i, \mathcal{D}_e) := N - F_{i_0+k}^1 \left[1 - \bar{h}_a \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right] - F_{i_0+k}^{-1}, \quad (52)$$

by the estimator [25]

$$\hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e) := N - \frac{NA_{i_0+k}}{C_{i_0+k}} \left\{ \hat{q}_x^1 P_{i_0+k}^1 \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right] + p_D \right\}. \quad (53)$$

The deviation $\Theta(\mathcal{D}_i, \mathcal{D}_e) - \hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e)$ is divided into two stochastic variables, $\Delta\Theta_1$ and $\Delta\Theta_2$;

$$\Theta(\mathcal{D}_i, \mathcal{D}_e) - \hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e) = \Delta\Theta_1 + \Delta\Theta_2, \quad (54)$$

where

$$\begin{aligned} \Delta\Theta_1 &:= -F_{i_0+k}^1 \left[1 - \bar{h}_a \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right] - F_{i_0+k}^{-1} \\ &\quad + \frac{N \left\{ \hat{q}_x^1 B_{i_0+k}^1 \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right] + A_{i_0+k} p_D \right\}}{C_{i_0+k}}, \end{aligned} \quad (55)$$

and

$$\Delta\Theta_2 := \frac{N \hat{q}_x^1 (A_{i_0+k} P_{i_0+k}^1 - B_{i_0+k}^1) \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right]}{C_{i_0+k}}. \quad (56)$$

Now, let us apply the Gaussian approximation to the variables $\Delta\Theta_1$ and $\Delta\Theta_2$. Since the mean of $\Delta\Theta_1$ is zero,

$$\sqrt{v_{i_0,x,y}}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu}) \Phi^{-1}(2^{-\delta_1}) \geq \Delta\Theta_1, \quad (57)$$

with probability $\geq 1 - 2^{-\delta_1}$. Here, $v_{i_0,x,y}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu})$ denotes the variance of $\Delta\Theta_1$, and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-\frac{y^2}{2}} dy \quad (58)$$

is the probabilistic distribution of the standard normal distribution. In Eq. (56), $A_{i_0+k} P_{i_0+k}^1 - B_{i_0+k}^1$ is a stochastic variable with the mean zero and the variance $A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)$. It follows that

$$\frac{N \hat{q}_x^1 \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right]}{C_{i_0+k}} \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_2}) \geq \Delta\Theta_2, \quad (59)$$

with probability $\geq 1 - 2^{-\delta_2}$. Consequently, taking the size of privacy amplification $m(\mathcal{D}_i, \mathcal{D}_e)$ in (11) as

$$\begin{aligned} m(\mathcal{D}_i, \mathcal{D}_e) &= \hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e) + \sqrt{v_{i_0,x,y}}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu}) \Phi^{-1}(2^{-\delta_1}) \\ &\quad + \frac{N \hat{q}_x^1 \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right]}{C_{i_0+k}} \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_2}) \\ &\quad + \delta_3, \end{aligned} \quad (60)$$

we have

$$P_{ph,av,\leftarrow}^P \leq 2^{-\delta_1} + 2^{-\delta_2} + 2^{-\delta_3}. \quad (61)$$

The quantity $m(\mathcal{D}_i, \mathcal{D}_e)$ [Eq. (60)] depends on \mathbf{q} , \mathbf{r} , and B_i^j , which can be approximated by $\mathbf{q}_{x,y,(B_i^j)}^{ML}$, $\mathbf{r}_{x,y,(B_i^j)}^{ML}$, and $A_i P_i^j$, respectively so that $m(\mathcal{D}_i, \mathcal{D}_e)$ can be written in terms of observed quantities. From above observation, we define the size

of privacy amplification for fixed x and y as

$$\begin{aligned}
& m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e) \\
& := \hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e) \\
& + \sqrt{v_{i_0;x,y}(\mathbf{q}_{x,y,(A_i P_i^j)}^{ML}, \mathbf{r}_{x,y,(A_i P_i^j)}^{ML}, A_i P_i^j, \mathbf{A}, \boldsymbol{\mu})} \Phi^{-1}(2^{-\delta_1}) \\
& + \frac{N \hat{q}_x^1 [1 - \bar{h}_a(\hat{r}_{x,y}^1)]}{C_{i_0+k}} \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_2}) \\
& + \delta_3,
\end{aligned} \tag{62}$$

which satisfies

$$m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e) \geq \Theta(\mathcal{D}_i, \mathcal{D}_e) + \delta_3 \tag{63}$$

with probability $1 - 2^{-\delta_1} - 2^{-\delta_2}$. It should be noted that the linear estimators \hat{q}_x^1 and $\hat{r}_{x,y}^1$ are used in the first term of the right-hand side of Eq. (62), while the maximally likelihood estimators are used in the second and third terms. This is because if linear estimators were used in the second and third terms in the right-hand side of Eq. (62), these terms would not well-defined since the linear estimators do not necessarily satisfy $0 \leq \hat{q}_x^j \leq 1 - p_D$ and $0 \leq \hat{r}_{x,y}^j \leq 1$.

Now, we take the worst case and define the size of privacy amplification $m_{i_0}(\mathcal{D}_i, \mathcal{D}_e)$ as

$$m_{i_0}(\mathcal{D}_i, \mathcal{D}_e) := \max_{0 \leq x \leq \sqrt{2}(1-p_D), 0 \leq y \leq 1} m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e), \tag{64}$$

then (61) holds for arbitrary x and y [26].

The size of privacy amplification for \times basis is given by $m_{i_0}(\tilde{\mathbf{A}}, \boldsymbol{\mu}, \tilde{p}_D, p_D; \tilde{\mathbf{C}}, \tilde{\mathbf{E}}, \tilde{\mathbf{H}})$ where $\tilde{A}_0 = A_0$, $\tilde{A}_i = A_{i+k}$, $\tilde{A}_{i+k} = A_i$, $\tilde{C}_0 = C_0$, $\tilde{C}_i = C_{i+k}$, $\tilde{C}_{i+k} = C_i$, $\tilde{E}_0 = E_0$, $\tilde{E}_i = E_{i+k}$, $\tilde{E}_{i+k} = E_i$, $\tilde{H}_i = H_{i+k}$, and $\tilde{H}_{i+k} = H_i$ for $i = 1, \dots, k$.

In the course of actual numerical computations, it often happens that x or y moves away from its defining region. To circumvent this difficulty, we use $\bar{h}_a(z)$ instead of $\bar{h}(z)$, which enables us to extend the accessible region of z to $(-\infty, \infty)$ and further to avoid the divergence of the derivative of \bar{h} . When a is small enough, $v_{i_0;x,y}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu})$ takes on a large value, while a is large, $\hat{\Theta}_{x,y}(\mathcal{D}_i, \mathcal{D}_e)$ takes on a large value in turn. The parameter a must be properly chosen taking into account such a trade-off behavior. The detail is shown in Appendix A.

Now, let us go back to Eq. (4). To ensure

$$I_{E,av,\leftarrow}^{\mathcal{P}} \leq 2^{-\delta}, \tag{65}$$

it is sufficient to choose $\delta_1 = \delta + \delta' + 1$ and $\delta_2 = \delta_3 = \delta + \delta' + 2$. Here, δ' is $\lceil \log_2 \bar{N} \rceil$. Since $\delta + \delta' \ll \bar{N}$, $\delta + \delta' + 1 + \bar{N} \lesssim 2^{\delta'}$. Thus,

$$\begin{aligned}
I_{E,av,\leftarrow}^{\mathcal{P}} & \leq P_{ph,av,\leftarrow}^{\mathcal{P}} (1 + \bar{N} - \log P_{ph,av,\leftarrow}^{\mathcal{P}}) \\
& \leq 2^{-\delta-\delta'} (1 + \bar{N} + \delta + \delta') \\
& \leq 2^{-\delta-\delta'} \cdot 2^{\delta'} = 2^{-\delta}.
\end{aligned} \tag{66}$$

In parallel with the above argument based on the Gaussian approximation, the large deviation type evaluation is also possible. By Cramér's theorem [27], $\text{Prob}\{|X - \mathbb{E}X| > cN\}$ goes

to zero exponentially when X obeys the N trials of a multinomial distribution or a multinomial hypergeometric distribution for an arbitrary constant $c > 0$. Hence, the probability satisfying the inequalities

$$\Delta\Theta_2 < c_1 N, \tag{67}$$

$\Delta F_{i_0+k}^1 = F_{i_0+k}^1 - \mathbb{E}(F_{i_0+k}^1) < c_2 N$, $\Delta G_{i_0+k}^1 = G_{i_0+k}^1 - \mathbb{E}(G_{i_0+k}^1) < c_3 N$, and $\Delta F_{i_0+k}^{-1} = F_{i_0+k}^{-1} - \mathbb{E}(F_{i_0+k}^{-1}) < c_4 N$ goes to zero exponentially for arbitrary constants $c_1, c_2, c_3, c_4 > 0$. For an arbitrary constant $c_5 > 0$, we choose $c_2, c_3, c_4 > 0$ such that if $\Delta F_{i_0+k}^1 < N c_2$, $\Delta G_{i_0+k}^1 < N c_3$, Δ , and $\Delta F_{i_0+k}^{-1} < N c_4$, then

$$\Delta\Theta_1 < c_5 N \tag{68}$$

is always satisfied. Thus, the probability satisfying (68) goes to zero exponentially. Now, we denote the exponential upper bounds of (67) and (68) by $2^{-d_1 N}$ and $2^{-d_2 N}$, respectively, and choose the size of sacrifice bits $m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e)$ as $\mathbb{E}\Theta + (c_1 + c_5 + c_6)N$ for a constant $c_6 > 0$. Then, applying (11), we obtain an exponential upper bound:

$$P_{ph,av,\leftarrow}^{\mathcal{P}} \leq 2^{-d_1 N} + 2^{-d_2 N} + 2^{-c_6 N}. \tag{69}$$

However, as is known in mathematical statistics, the above exponential evaluation does not yield a tighter upper bound of $\text{Prob}\{|X - \mathbb{E}X| > cN\}$ as the Gaussian approximation does. Hence, in the finite-length code, the Gaussian approximation gives a tighter upper bound of $P_{ph,av,\leftarrow}^{\mathcal{P}}$ than the above exponential upper bound. This is the reason why we have employed the Gaussian approximation.

VII. COMPUTATION OF $v_{i_0;x,y}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu})$

The variance $v_{i_0;x,y}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu})$ for a given stochastic variable B_i^j and under constraints $\frac{q^{k+1} + q^{2k+1}}{\sqrt{2}} = x$ and $r^{k+1} = y$ is given by

$$v_{i_0;x,y}(\mathbf{q}, \mathbf{r}, B_i^j, \mathbf{A}, \boldsymbol{\mu}) = \mathbb{E}\Delta\Theta_1^2, \tag{70}$$

where

$$\begin{aligned}
& \Delta\Theta_1 \\
& \cong - \left[1 - \bar{h}_a(r^1) \right] \Delta F_{i_0+k}^1 + \bar{h}'_a(r^1) \Delta' G_{i_0+k}^1 - \Delta F_{i_0+k}^{-1} \\
& - \frac{N \left\{ q^1 B_{i_0+k}^1 \left[1 - \bar{h}_a(r^1) \right] + A_{i_0+k} p_D \right\}}{(\mathbb{E}C_{i_0+k})^2} \Delta C_{i_0+k} \\
& + \frac{N B_{i_0+k}^1 \left[1 - \bar{h}_a(r^1) \right]}{\mathbb{E}C_{i_0+k}} \Delta \hat{q}_x^1 - \frac{N}{\mathbb{E}C_{i_0+k}} q^1 B_{i_0+k}^1 \bar{h}'_a(r^1) \Delta \hat{r}_{x,y}^1,
\end{aligned} \tag{71}$$

with $\Delta X = X - \mathbb{E}X$ in the right-hand side of Eq. (71). Here, stochastic variables, $\Delta F_{i_0+k}^1$, $\Delta F_{i_0+k}^{-1}$, $\Delta \hat{q}_x^1$, and $\Delta \hat{r}_{x,y}^1$ are computed as

$$\begin{aligned}
\Delta F_{i_0+k}^j & \cong \frac{2N}{\mathbb{E}C_{i_0+k}} \Delta E_{i_0+k}^j - \frac{2N q^j B_{i_0+k}^j}{(\mathbb{E}C_{i_0+k})^2} \sum_{j'=-1}^{2k+1} \Delta E_{i_0+k}^{j'} \\
& + \Delta' F_{i_0+k}^j,
\end{aligned} \tag{72}$$

$$\Delta\hat{q}_x^1 = \sum_{i=1}^{2k} (\bar{P}_{2k \times 2k}^{-1})_1^i \left(\frac{\Delta C_i}{A_i} - \bar{P}_i^0 \frac{\Delta C_0}{A_i} \right), \quad (73) \quad \text{and}$$

$$\begin{aligned} \Delta\hat{q}_{x,y}^1 &\cong -\frac{1}{(1-2p_S)(\hat{q}_x^1)^2} \sum_{i=1}^k (P_{k \times k}^{-1})_1^i \left\{ \mathbb{E}(s_i p_i) - \frac{1}{2} \left[P_i^0 \left(\frac{\mathbb{E}C_0}{A_0} - p_D \right) + p_D \right] - P_i^{k+1} q^{k+1} y \right\} \Delta\hat{q}_x^1 \\ &\quad + \frac{1}{(1-2p_S)\hat{q}_x^1} \sum_{i=1}^k (P_{k \times k}^{-1})_1^i \left(\Delta(s_i p_i) - \frac{1}{2} P_i^0 \frac{\Delta C_0}{A_0} - P_i^{k+1} y \Delta\hat{q}_x^{k+1} \right). \end{aligned} \quad (74)$$

In Eq. (74), $\Delta\hat{q}_x^{k+1}$ and $\Delta(s_i p_i)$ are given by

$$\Delta\hat{q}_x^{k+1} = \frac{1}{\sqrt{2}} \sum_{i=1}^{2k} (\bar{P}_{2k \times 2k}^{-1})_{k+1}^i \left(\frac{\Delta C_i}{A_i} - \bar{P}_i^0 \frac{\Delta C_0}{A_i} \right), \quad (75)$$

$$\Delta(s_i p_i) \cong \frac{2}{A_i} \Delta H_i - \frac{2\mathbb{E}H_i}{A_i \mathbb{E}E_i} \Delta' E_i, \quad (76)$$

for $i = 1, \dots, k; i \neq i_0$ and

$$\begin{aligned} \Delta(s_{i_0} p_{i_0}) &\cong \frac{\mathbb{E}C_{i_0}}{A_{i_0}(\mathbb{E}E_{i_0} - N)} \Delta H_{i_0} - \frac{\mathbb{E}C_{i_0} \mathbb{E}H_{i_0}}{A_{i_0}(\mathbb{E}E_{i_0} - N)^2} \Delta' E_{i_0} \\ &\quad - \frac{N \mathbb{E}H_{i_0}}{A_{i_0}(\mathbb{E}E_{i_0} - N)^2} \Delta C_{i_0}. \end{aligned} \quad (77)$$

Deviations ΔH_i in Eqs. (76) and (77) are computed as

$$\begin{aligned} \Delta H_i &= \sum_{j=-1}^{k+1} \frac{\tilde{r}^j}{2} \Delta C_i^j + \sum_{j=-1}^{k+1} \tilde{r}^j \Delta' E_i^j + (1-2p_S) \Delta' G_i^1 \\ &\quad + \sum_{j=2}^{k+1} \Delta' G_i^j + \Delta' H_i, \end{aligned} \quad (78)$$

for $i = 1, \dots, k; i \neq i_0$. and

$$\begin{aligned} \Delta H_{i_0} &= \sum_{j=-1}^{k+1} \tilde{r}^j \left[\left(1 - \frac{2N}{\mathbb{E}C_{i_0}} \right) \Delta E_{i_0}^j \right. \\ &\quad \left. + \frac{2N \mathbb{E}C_{i_0}^j}{(\mathbb{E}C_{i_0})^2} \sum_{j'=-1}^{2k+1} \Delta E_{i_0}^{j'} + \Delta' F_{i_0}^j \right] \\ &\quad + (1-2p_S) \Delta' G_{i_0}^1 + \sum_{j=2}^{k+1} \Delta' G_{i_0}^j + \Delta' H_{i_0}, \end{aligned} \quad (79)$$

where $\tilde{r}^{-1} = \tilde{r}^0 = \frac{1}{2}$, $\tilde{r}^1 = r^1 + (1-2r^1)p_S$, and $\tilde{r}^j = r^j$ for $j \geq 2$. Note that $\Delta E_i^j = \frac{1}{2} \Delta' C_i^j + \Delta' E_i^j$ in Eqs. (72) and (79). Equation (70) with subsequent equations in this section and (co)variances given in Sec. V yields the explicit form of $v_{i_0,x,y}$.

If \mathbf{q} and \mathbf{r} coincide with the respective values estimated from the observed quantities $\mathcal{D}_e = (\mathbf{C}, \mathbf{E}, \mathbf{H})$, $\mathbb{E}C_i$, $\mathbb{E}E_{i_0}$, and $\mathbb{E}H_i$ in Eqs. (71), (72), (76), and (77) are, respectively, equal to the observed quantities, C_i , E_{i_0} , and H_i .

TABLE I: Parameters for numerical simulation.

a_1 (db/km)	a_0 (db)	η_{det}
0.17	5.0	0.1
p_D	p_S, \tilde{p}_S	p_0
4.0×10^{-7}	3%	4.0×10^{-7}

VIII. NUMERICAL ANALYSIS

In this section we show the results of numerical simulation to reveal performances of our protocol and to know the optimal values of parameters such as intensities in our protocol. For numerical simulation, we use $k+1 = 4$ different intensities including vacuum because $k = 3$ is required for at least good estimation of the probability that multi-photon is actually sent by Alice [22].

A. Parameters

Let a_1 (db/km) be the fiber loss, a_0 (db) be the receiver loss, η_{det} be the efficiency of the detector, and L (km) be the transmission distance. A detection probability p of the pulse μ_i can be represented by

$$p = 1 - e^{-a\mu_i} + p_0, \quad (80)$$

where

$$\alpha = \eta_{det} \cdot 10^{-\frac{a_1 L + a_0}{10}}. \quad (81)$$

Listed in Table I are parameters used for our numerical simulation, all of which are experimental values in the long distance experiment [28], and a_1 is the lowest loss value in commercially available optical fibers [29]. We assume that the detection probability p_0 of vacuum state equals to dark count rate p_D . In this setting, the detection probability of the pulse $\mu = 0.5$ at $L = 20.0(100.0)$ (km) is $7.2 \times 10^{-3}(3.2 \times 10^{-3})$ and the error probability is 3.00%(3.06%), respectively.

We also fix the security parameter δ in order to satisfy the average of Eve's information $I_{E,av,\leftarrow}^P \leq 2^{-9}$ in Eq. (65). It

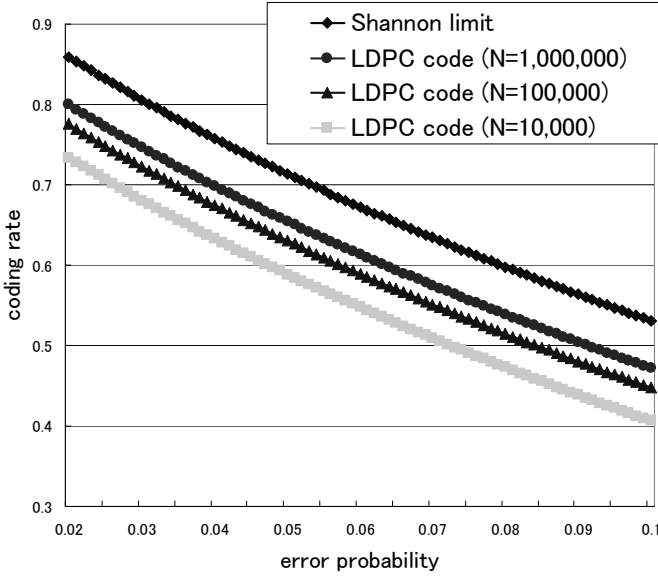


FIG. 1: Coding rates of the LDPC codes used for numerical simulation.

is sufficient to set $\delta_1 = \delta_3 = 9 + \lceil \log_2(\bar{N}) \rceil + 2$, and $\delta_2 = 9 + \lceil \log_2(\bar{N}) \rceil + 1$.

For the error correction with the finite code length, we use the LDPC (Low Density Parity Check) code [30, 31] which is known to be one of the best classical error correcting codes and the performance can asymptotically achieve the Shannon limit. Figure 1 shows coding rates of our LDPC codes. The coding rate with $N = 1.0 \times 10^6$ is about 0.75 when the error probability is 3%.

Suppose that intensities μ_i and sending probabilities \tilde{p}_i take discrete values 0.05, 0.10, ..., 1.00 for our numerical simulation because computation of $m_{i_0; x, y}(\mathcal{D}_i, \mathcal{D}_e)$ (62) needs a constrained non-linear optimization and it is therefore rather hard and time-consuming task.

B. Performance

We first show key generation rates of our protocol with respect to the transmission distance. Figure 2 shows the optimal key generation rates per pulse sent by Alice when the code length N equals to 1.0×10^4 , 1.0×10^5 , and 1.0×10^6 , keeping the average of Eve's information less than 2^{-9} . For comparison, the asymptotic rate is also added, which is calculated from the asymptotic rate formula with three different intensities including vacuum [22]. Our decoy state QKD enables Alice and Bob to share the final secret key securely up to 150 (km) while the secure secret key can be shared up to 250 (km) in the asymptotic case. Although there is a huge gap between the maximum transmission distances of the finite and asymptotic case, such small key generation rate with the finite code length, in other words the large number of sacrifice bits, is needed for keeping the average of Eve's information less than 2^{-9} incorporating the statistical fluctuations. The main reason

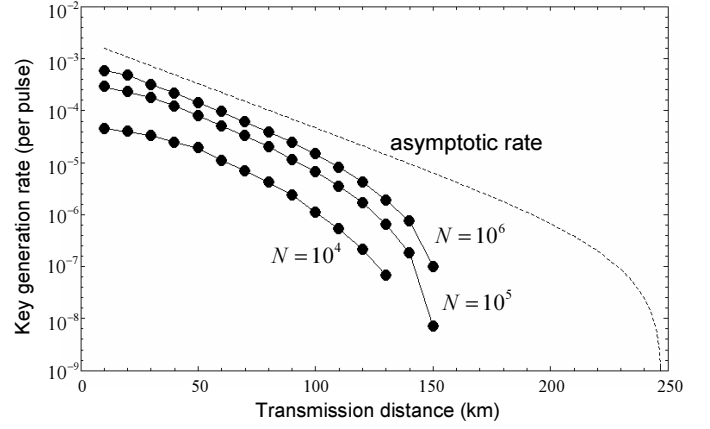


FIG. 2: Performance of our decoy state QKD. The secure secret key can be shared up to 150 km when $N = 1.0 \times 10^6$.

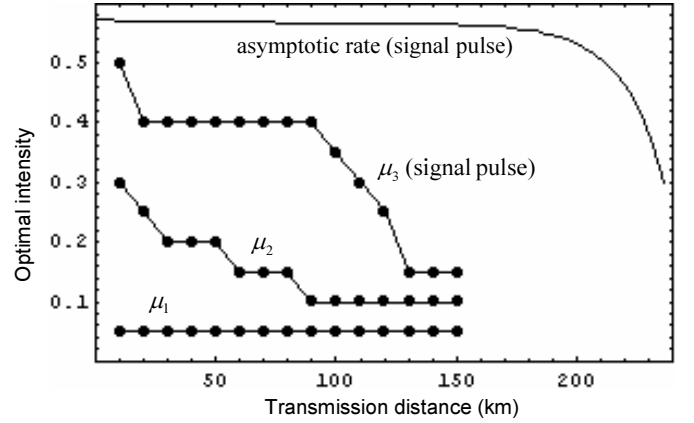


FIG. 3: Optimal intensities with $N = 1.0 \times 10^6$ and the asymptotic case corresponding to the key generation rates in Figure 2.

why the more key generation rate is obtained with larger code size is that Eve's information can be estimated better with large size statistics, such that the larger number of sampling given by the check bits can be used for estimation as well as the finite error correcting coding rate in Figure 1.

We then show the optimal intensities corresponding to the rates in Figure 2. The three optimal different intensities μ_i with $N = 1.0 \times 10^6$ as well as the signal intensity in the asymptotic case are shown in Figure 3. The optimal signal intensities decrease as the transmission distance is longer. This tendency is easy to understand because the probability that the multi-photon is emitted cannot be negligible for estimation of the quantum channel at the long transmission distance. At the maximum transmission distance 150 (km), The optimal intensity μ_3 becomes 0.15 which is the minimum value of the signal intensity available for our numerical simulation because intensities μ_i take discrete values by 0.05 and $0 < \mu_1 < \mu_2 < \mu_3$. It is quite smaller than the intensity 0.3 of the asymptotic case at the maximal transmission distance 250 (km). This is be-

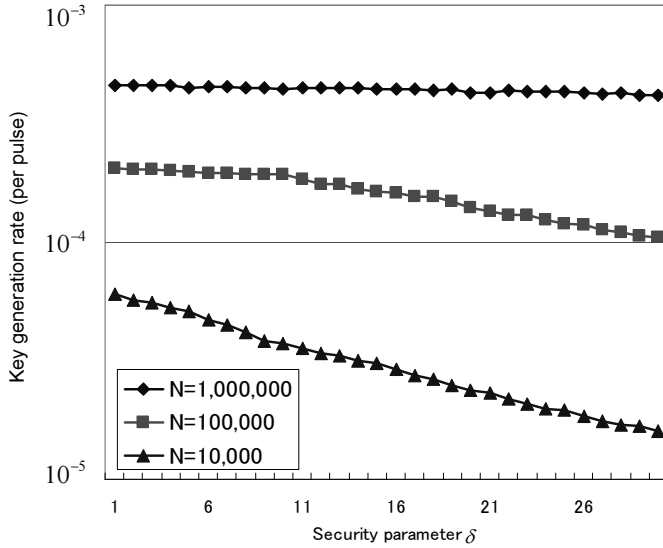


FIG. 4: Key generation rates with respect to the security parameter δ when the code length $N = 10^4, 10^5$, and 10^6 .

cause considering statistical fluctuations including estimation errors causes the worse estimations of $m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e)$ and such random variables as \mathbf{J} than the true values without estimation errors. It turns out that the probability that the multi-photon is sent including statistical fluctuations becomes more dominant factor in our protocol.

We finally examine the performance of our protocol with respect to the security parameter δ . Figure 4 shows the key generation rates at the transmission distance 20 (km). The rates when $N = 1.0 \times 10^6$ hardly decrease at all while the rates when $N = 1.0 \times 10^4$ and 1.0×10^5 become smaller as δ is bigger. This is because the larger number of sampling when N is large enough makes the variance $v_{i_0;x,y}$ in Eq (62) smaller. Therefore the size of sacrifice bits $m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e)$ is affected little by δ_1 , that is δ , and the final secret key can be shared as securely as you want when $N = 1.0 \times 10^6$.

IX. CONCLUSIONS AND FUTURE PERSPECTIVES

In conclusion, we have derived a formula for the size the final secret key in the finite code length decoy state QKD with arbitrary number of decoy states of different intensities incorporating the finite statistics [32].

We have utilized the central limit theorem and thereby neglected the higher-order terms in the formula obtained here. Furthermore, there is room to improve the derivation of the size of privacy amplification. These points will be further pursued in the future. Finally, we have assumed that the error probability of the signal generation at the sending port is unknown. If this is not the case, the generation rate of the final secret key would be improved [33]. Such an improvement of arguments in Sec. VI would be also one of the future problems.

APPENDIX A: CHOICE OF PARAMETER a

In Sec. VI we have used $\bar{h}_a(x)$ instead of $\bar{h}(x)$ to circumvent the singularity of $\bar{h}(x)$ when $x \rightarrow 0$. If x is so close to 0 that $\bar{h}'_a(x) \gg \bar{h}_a(x)$, the leading term of $\Delta\Theta_1$ is well approximated by

$$\bar{h}'_a(\hat{r}_{x,y}^1) \left(\Delta' G_{i_0+k}^1 - \frac{N}{\mathbb{E}C_{i_0+k}} q^1 B_{i_0+k}^1 \Delta \hat{r}_{x,y}^1 \right).$$

Hence, denoting the variance of

$$\Delta' G_{i_0+k}^1 - \frac{N}{\mathbb{E}C_{i_0+k}} q^1 B_{i_0+k}^1 \Delta \hat{r}_{x,y}^1$$

by V , we have

$$\begin{aligned} m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e) &\cong N - \frac{NA_{i_0+k}(\hat{q}_x^1 P_{i_0+k}^1 + p_D)}{C_{i_0+k}} + \bar{h}'_a(\hat{r}_{x,y}^1) \sqrt{V} \Phi^{-1}(2^{-\delta_1}) \\ &+ \frac{N\hat{q}_x^1}{C_{i_0+k}} \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_2}) \\ &+ \bar{h}_a(\hat{r}_{x,y}^1) \frac{N\hat{q}_x^1}{C_{i_0+k}} \left[A_{i_0+k} P_{i_0+k}^1 \right. \\ &\left. - \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_2}) \right] + \delta_3. \end{aligned} \quad (\text{A1})$$

Our task is to choose the parameter a that minimizes $m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e)$; the problem is reduced to the minimization of $f(a) = \bar{h}_a(\hat{r}_{x,y}^1)S + \bar{h}'_a(\hat{r}_{x,y}^1)T$ with respect to a , where

$$f'(a) = \begin{cases} (S(r-a) + T)h''(a) & \hat{r}_{x,y}^1 \geq a, \\ 0 & \hat{r}_{x,y}^1 < a, \end{cases} \quad (\text{A2})$$

$$S := \frac{N\hat{q}_x^1}{C_{i_0+k}} \left[A_{i_0+k} P_{i_0+k}^1 - \sqrt{A_{i_0+k} P_{i_0+k}^1 (1 - P_{i_0+k}^1)} \Phi^{-1}(2^{-\delta_1}) \right], \quad (\text{A3})$$

and

$$T := \sqrt{V} \Phi^{-1}(2^{-\delta_2}). \quad (\text{A4})$$

Since $h''(a) \leq 0$, the minimal $f(a)$ is achieved when $a = \hat{r}_{x,y}^1 + T/S$. The values of S and T are actually unknown and vary from time to time. However, if we can expect almost constant values for S and T , which can be measured in advance, a favorable choice of a is $a = \hat{r}_{x,y}^1 + T/S$.

APPENDIX B: COMPUTATION OF SACRIFICE KEY SIZE (FORWARD CASE)

To compute the size of sacrifice bits in the case of forward error correction, we firstly change the definitions of random variables in Sec. III since J^3 in Eq. (10) cannot be expressed in terms of them with their original definitions. Major alteration concerns the definition of the subscript j : For

$j = 0, 2, \dots, 2k + 1$, q^j are changed to represent the detection ratio *including* the detector dark counts and q^{-1} is changed to stand for the dark count ratio given that the emitted state is a single photon state (q^1 is left unchanged). Namely, for $j = 0, 2, \dots, 2k + 1$, C_i^j now denote the numbers of pulses detected normally as well as by dark counts given that the emitted state is ρ_j and C_i^{-1} now stands for the number of dark counts given that the emitted state is ρ_1 . Alongside the meaning of r^j is changed for $j = 2, \dots, k + 1$. According to these alterations, the definition of random variables E_i^j , F_i^j , G_i^j , and H_i are also subject to modification. Almost all of equations in Sec. IV and Sec. V are left unchanged except that (i) p_D should read $p_D P_i^1$. (ii) the range of q^j ($j = 2, \dots, 2k$) should be changed to $0 \leq q^j \leq 1$. (iii) the range of x should be changed to $[1, \sqrt{2}]$. and (iv) Eqs. (22), (24), and (25) should read, respectively,

$$\hat{q}_x^j = \begin{cases} p_0 & j = 0 \\ \xi^j & j = 1, \dots, k, k + 2, \dots, 2k \\ \frac{1}{\sqrt{2}}(x + \xi^{k+1}) & j = k + 1 \\ \frac{1}{\sqrt{2}}(x - \xi^{k+1}) & j = 2k + 1, \end{cases} \quad (\text{B1})$$

$$\zeta^j := \frac{1}{\hat{q}_x^j} \sum_{i=1}^k (P_{k \times k}^{-1})_j^i \times \left[s_i p_i - \frac{1}{2} (P_i^0 p_0 + P_i^1 p_D) - P_i^{k+1} \hat{q}_x^{k+1} y \right], \quad (\text{B2})$$

and

$$\xi^j := \sum_{i=1}^{2k} (\bar{P}_{2k \times 2k}^{-1})_j^i \left(p_i - \bar{P}_i^0 p_0 - \bar{P}_i^1 p_D - x \bar{P}_i^{2k+1} \right). \quad (\text{B3})$$

The upper bound for the average phase error is now given by

$$\frac{1}{2} \left[m - N + F_{i_0+k}^1 \left(1 - \bar{h} \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right) + F_{i_0+k}^0 \right]_+,$$

Here, we use

$$N - \frac{N A_{i_0+k}}{C_{i_0+k}} \left\{ \hat{q}_x^1 P_{i_0+k}^1 \left[1 - \bar{h}_a(\hat{r}_{x,y}^1) \right] + \frac{C_0}{A_0} \right\}$$

as the estimation of

$$N - F_{i_0+k}^1 \left[1 - \bar{h} \left(\frac{G_{i_0+k}^1}{F_{i_0+k}^1} \right) \right] - F_{i_0+k}^0.$$

The subsequent argument is parallel to that in Sec. VI.

-
- [1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India, 1984* (IEEE, New York, 1984), p. 175.
- [2] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Rev. Mod. Phys.* **74**, 145 (2002).
- [3] D. Mayers, in *Advances in Cryptology — Proceedings of Crypto '96*; Lecture Notes in Computer Science, **1109**, 343 (1996); *J. ACM* **48**, 351 (2001).
- [4] H.-K. Lo and H. F. Chau, *Science* **283**, 2050 (1999); P. W. Shor and J. Preskill, *Phys. Rev. Lett.* **85** 411 (2000); H. Inamori, N. Lütkenhaus, and D. Mayers, quant-ph/0107017.
- [5] H. Inamori, N. Lütkenhaus, and D. Mayers, quant-ph/0107017.
- [6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, United Kingdom, 2000).
- [7] M. Hayashi, *Quantum Information: An Introduction* (Springer-Verlag, Berlin, 2006).
- [8] B. Huttner, N. Imoto, N. Gisin, and T. Mor, *Phys. Rev. A* **51**, 1863 (1995); G. Brassard, N. Lütkenhaus, T. Mor, and B. Sanders, *Phys. Rev. Lett.* **85**, 1330 (2000); N. Lütkenhaus and M. Jähma, *New J. Phys.* **4** 44 (2002).
- [9] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Quantum Inf. Comput.* **4**, 325 (2004).
- [10] W.-Y. Hwang, *Phys. Rev. Lett.* **91** 057901 (2003).
- [11] X.-B. Wang, *Phys. Rev. A* **72** 012322 (2005); *Phys. Rev. Lett.* **94** 230503 (2005).
- [12] H.-K. Lo, X. Ma, and K. Chen, *Phys. Rev. Lett.* **94** 230504 (2005).
- [13] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, *Phys. Rev. A* **72**, 012326 (2005); X. Ma, C.-Z. F. Fung, F. Dupuis, K. Chen, K. Tamaki, and H.-K. Lo, *ibid.* **74** 032330 (2006).
- [14] Y. Zhao, B. Qi, X. Ma, H.-K. Lo, and L. Qian, *Phys. Rev. Lett.* **96** 070502 (2006).
- [15] D. Rosenberg, J. W. Harrington, P. R. Rice, P. A. Hiskett, C. G. Peterson, R. J. Hughes, A. E. Lita, S. W. Nam, and J. E. Nordholt, *Phys. Rev. Lett.* **98** 010503 (2007).
- [16] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter, *Phys. Rev. Lett.* **98** 010504 (2007).
- [17] C.-Z. Peng, J. Zhang, D. Yang, W.-B. Gao, H.-X. Ma, H. Yin, H.-P. Zeng, T. Yang, X.-W. Wang, and J.-W. Pan, *Phys. Rev. Lett.* **98** 010505 (2007).
- [18] Z. L. Yuan, A. W. Sharpe, and A. J. Shields, *Appl. Phys. Lett.* **90** 011118 (2007).
- [19] M. Hayashi, *Phys. Rev. A* **74**, 022307 (2006).
- [20] J. W. Harrington, J. M. Ettinger, R. J. Hughes, and J. E. Nordholt, quant-ph/0503002.
- [21] M. Hayashi, "Upper bounds of security parameters in finite-length code with decoy method," *Phys. Rev. A*, to appear; quant-ph/0702250.
- [22] M. Hayashi, "Asymptotic key generation rates with phase-randomized coherent light by decoy method," quant-ph/0702251.
- [23] More generally, Eve can choose $q(\mathbf{B})^j$ stochastically according

to the value of \mathbf{B} . Such an operation can be regarded as the probabilistic mixture of the cases when $q(\mathbf{B})^j$ is chosen deterministically. However, it is sufficient to consider Eve's operation in an extremal point for the security arguments [21].

- [24] The state ρ_{-1} is regarded as the pulse detected by dark counts.
 [25] We adopt the linear estimator here. Otherwise if we adopted the maximal likelihood estimator, the computation of variance $v_{i_0;x,y}$ would be intractable.
 [26] If the light source is unstable, for example, the values of B_i^j may move somewhat away from the ideal values $A_i P_i^j$ [34, 35]. In this case, $m_{i_0;x,y}(\mathcal{D}_i, \mathcal{D}_e)$ should be replaced by

$$\begin{aligned} & m'_{i_0;x,y}(B_i^j, \mathcal{D}_i, \mathcal{D}_e) \\ &= N + \frac{N \left\{ \hat{q}_x^1 B_{i_0+k}^1 \left[\bar{h}_a(\hat{r}_{x,y}^1) - 1 \right] - A_{i_0+k} p_D \right\}}{C_{i_0+k}} \\ &+ \sqrt{v_{i_0;x,y}(\mathbf{q}_{x,y,(B_i^j)}^{ML}, \mathbf{r}_{x,y,(B_i^j)}^{ML}, B_i^j, \mathbf{A}, \boldsymbol{\mu}) \Phi^{-1}(2^{-\delta_1})} + \delta_3. \end{aligned}$$

Suppose \mathbf{B} belongs to the set \mathcal{X} with probability $1 - \epsilon$. Then, if we define the size of privacy amplification $m'_{i_0}(\mathcal{D}_i, \mathcal{D}_e)$ as

$$m'_{i_0}(\mathcal{D}_i, \mathcal{D}_e) = \max_{\mathbf{B} \in \mathcal{X}, 0 \leq x \leq \sqrt{2}(1-p_D), 0 \leq y \leq 1} m'_{i_0;x,y}(B_i^j, \mathcal{D}_i, \mathcal{D}_e),$$

we have from (11), $P_{ph,av,\leftarrow}^{\mathcal{P}} \leq \epsilon + 2^{-\delta_1} + 2^{-\delta_3}$.

- [27] J. A. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation* (Wiley, New York, 1990).
 [28] T. Kimura, Y. Nambu, T. Hatanaka, A. Tomita, H. Kosaka, and K. Nakamura, *Japanese Journal of Applied Physics* **43**, 9A/B, p.p. 1217-1219 (2004).
 [29] http://www.corning.com/opticalfiber/products_applications/products/smf_28_ULL.aspx, <http://www.sei.co.jp/fbr-opt-eng/submarine/zplus/pdf/zplus1.pdf>.
 [30] R. G. Gallager, *Low-Density Parity-Check Codes* (MIT Press, Cambridge, MA, 1963).
 [31] D. J. C. MacKay, in *Proceedings of IEEE Transactions on Information Theory*, **45** (2), p.p. 399-431 (1999).
 [32] Since any photon states can be expanded in terms of coherent states, the formalism presented in this paper is still valid for any distributions of photons from the light source by changing the generating probability matrices [Eq. (17)] appropriately.
 [33] R. Renner, N. Gisin, and B. Kraus, *Phys. Rev. A* **72**, 012332 (2005).
 [34] X.-B. Wang, *Phys. Rev. A* **75**, 012301 (2007).
 [35] X.-B. Wang, C.-Z. Peng, and J.-W. Pan, *Appl. Phys. Lett.* **90**, 031110 (2007).